



3D-QSAR analysis of metallocene-based catalysts used in ethylene polymerisation

V. Cruz^a, J. Ramos^b, A. Muñoz-Escalona^b, P. Lafuente^c, B. Peña^c, J. Martínez-Salazar^{b,*}

^aDepartamento de Física Macromolecular, Instituto de Estructura de la Materia C.S.I.C., Serrano 113b, 28006 Madrid, Spain

^bCentro Técnico de Informática C.S.I.C., Pinar 19, 28006 Madrid, Spain

^cCentro de Investigación REPSOL-YPF, Ctra. de Móstoles, salida 18, Móstoles, Madrid, Spain

Received 13 November 2003; received in revised form 19 December 2003; accepted 22 December 2003

Abstract

Single site olefin polymerisation catalysts are suitable candidates for modelling purposes. Their well-defined structure and the almost complete elucidation of their polymerisation mechanisms, make these organometallic complexes ideal for studies based on quantitative structure-activity relationships (QSAR). Although the QSAR technique is extensively used in drug design, there are very few reports on its application to metallocene-based polymerisation catalysis. This probably has something to do with the difficulties inherent in controlling experimental conditions during the polymerisation process. In the present study, we obtained ethylene polymerisation data using a number of zirconocene catalysts under carefully controlled experimental conditions, i.e. keeping all polymerisation variables constant except catalyst structure. The catalytic activity and molecular weight of the resulting polyethylenes were experimentally determined. We then applied 3D-QSAR methodology to explain the experimental data in terms of three-dimensional (3D) field descriptors related to the structure of the metallocene catalyst. Our results provide useful correlations between experimental ethylene polymerisation activities and the steric, LUMO and local softness fields of the catalysts. The molecular weights calculated from comparative molecular field analysis (CoMFA) models including LUMO and local softness fields correlate well with the experimental ones. The predictive capacity of the models was also tested. Based on the proposed models, steric and electronic factors affecting polymerisation performance are discussed.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Metallocene catalysts; Quantitative structure-activity relationships; DFT calculations

1. Introduction

Metallocenes are extensively used as catalysts for olefin polymerisation in place of the traditional Ziegler-Natta catalysts. Many experimental attempts have been made to correlate catalyst structure and polymerisation conditions with catalytic activity. Among other factors, the temperature, solvent, monomer pressure, the nature of the catalyst and the catalyst/co-catalyst ratio have been found to considerably affect polymerisation activity. If all these conditions are kept constant, the differences in polymerisation activity observed for a series of metallocene catalysts can be directly related to the molecular structure of the organometallic complexes. In a recent review, Alt et al. [1] reported they had synthesized over 650 metallocene and half-sandwich catalyst precursors to test their catalytic

potential in olefin polymerisation. These authors concluded that there are many factors involved in polymerisation activity, and that qualitative trends could be appropriately inferred through molecular modelling or experiments. Systematic experiments performed with different catalysts under fixed polymerisation conditions are thus highly suitable for quantitative structure-activity relationship (QSAR) studies.

The QSAR approach is a rational method of drug design [2]; the synthesis and screening of hundreds of potential drug candidates being extremely expensive and laborious. The success of QSAR in medicinal chemistry stems from the fact that the single most significant event in pharmacology is the interaction between a drug molecule and its receptor [3]. Although there is a vast amount of scientific literature on QSAR work in medicinal chemistry, very few authors have applied the concept to the polymerisation activity of metallocene catalysts. One of the reasons

* Corresponding author. Tel.: +34-915-901-618; fax: +34-915-855-413.
E-mail address: jmsalazar@iem.cfmac.csic.es (J. Martínez-Salazar).

proposed for this lack of systematic QSAR studies is the difficulty in obtaining accurate experimental polymerisation activity data. Möring and Coville [4] reported a quantitative relationship between the catalytic activity of $(\text{CpR})_2\text{ZrCl}_2$ (where R is a cyclopentadienyl monoalkyl ring substituent) and steric and electronic descriptors. These authors noted increased activity of the catalytic system with increasing size of the R ligands, as well as with the increasing electron-donating capacity of the substituents. Yao et al. [5] performed a QSAR analysis on the influence of the catalyst/co-catalyst ion pair structure on polymerisation activity. They used molecular mechanics and dynamics to describe the molecular geometries of the ion pair. Activity differences in their model were explained by the space between the catalyst and co-catalyst systems. More recently, Linnolahti et al. [6] undertook an extensive qualitative study to evaluate the effect of the ligand structure of zirconocene catalytic systems on the accessibility and relative stability of the active reaction centres. By comparing experimental and theoretical results, these authors demonstrated a qualitative correlation between active reaction centre accessibilities and polymerisation activity.

In this paper, we report a first attempt at using three-dimensional QSAR (3D-QSAR) methodology to predict the behaviour of several metallocene catalysts in ethylene polymerisation. The idea underlying 3D-QSAR analysis is that differences in a target property (in this case polymerisation activity and polymer molecular weight) are related to differences in the shapes and intensities of the molecular fields describing the tested molecules [7]. 3D-QSAR methods identify spatial regions of a given molecule that match the binding site of the receptor [8]. The methodology most widely applied in this field is the so-called comparative molecular field analysis (CoMFA) [9], which is based on identifying critical local differences by sampling field intensities on a Cartesian lattice. Further, the 3D information is subsequently de-folded into a 2D table containing hundreds of descriptors for each molecule. This table is finally analysed by partial least squares (PLS) statistical modelling, or PLS [10], to find out which descriptors, and therefore grid locations, best explain variation in a measured property. Since the original 3D descriptors calculated at each grid point are not independent variables they cannot be used by themselves in the regression analysis. To overcome this, PLS extracts components, known as ‘latent variables’. These are underlying patterns in the 3D descriptors with the additional property that they are mutually orthogonal or independent and therefore suitable for statistical analysis. The procedure used in PLS to assess the predictive ability of a 3D-QSAR model is the cross-validation technique [11], in which regression equations are validated according to their predictive power. The optimal number of components in the 3D-QSAR equations is obtained from the cross-validation process. Final PLS models are derived by fitting all the experimental values to the equation containing the optimal number of components calculated in the previous step.

In a similar manner to the usual practice of 3D-QSAR in drug design [12], we tried to model the different metallocene catalysts taken as the ‘receptor’, and the ethylene monomer as the interacting ‘ligand’ in each case. Herein, we report the results of a first attempt to correlate experimental polymerisation activity and polymer molecular weights with 3D fields associated with catalyst structure. Our results provide relevant information on the structural and electronic characteristics that affect polymerisation activity and polymer molecular weight. They also indicate that the 3D-QSAR method may serve to identify the main structural features required to develop better metallocene catalysts for olefin polymerisation.

2. Methods

2.1. Experimental data

Table 1 shows the experimental activities and molecular weights of seven metallocene catalysts that are illustrated in Scheme 1. Polymerisation reactions were performed in a Büchi reactor using *n*-heptane as solvent under the following conditions: 70 °C polymerisation temperature, Al/M ratio = 1500, catalyst concentration = 0.8 μmol and ethylene pressure = 4 atm. Molecular weights were determined by size exclusion gel permeation chromatography (GPC) in 1,2,4-trichlorobenzene (TCB) solvent at a flow rate of 1.0 ml/min at 145° using a Water GPC 150C instrument equipped with three mixed bed columns and refractive index and viscometer detectors.

2.2. Molecular modelling

The active species of the catalysts is a cationic organometallic complex with a vacant coordination site that can interact with the co-catalyst where polymerisation takes place. Thus, the calculations are based on the cationic species rather than in the precursor metallocene itself. Furthermore, although the co-catalyst is not taken into account explicitly in the present work, the possible effect on the activity and molecular weight is assumed to be the same in all cases. This assumption is valid as far as the same co-catalyst and the same Al/Zr ratio were used for all the polymerisation reactions. For each organometallic cationic species, geometry optimisation at B3LYP [13]/LANL2DZ [14] was performed using the Gaussian98 package [15]. The following 3D fields were evaluated in Cartesian grids: electron densities, electrostatic potential, HOMO and LUMO molecular orbitals. By fitting the electrostatic potential to nuclear positions according to the CHELPG [16] scheme, we also calculated electrostatic charges. In the CoMFA method [9], steric and electrostatic 3D fields are calculated through the interaction between each catalyst and a probe atom. The probe atom should have specific charge and steric properties to evaluate the interaction energy at

Table 1
Catalyst experimental data

Number ^a	Activity ^b	Molecular weight ^c	CpZrCp angle ^d	ASA ^e	Global softness ^f
1	19.5	145	127.5	5.1	9.2
2	11.9	320	136.0	5.7	7.3
3	8.5	270	132.8	7.5	6.6
4	5.0	195	127.9	5.3	9.1
5	4.0	180	128.8	6.2	7.4
6	2.0	75	125.0	7.2	6.7
7	1.0	35	125.7	5.7	7.3

^a See Scheme 1 for labelling.

^b Catalyst activity in kg PE (mol Zr × h × [Ethylene])⁻¹ × 10⁻³.

^c *M_w* in kg/mol.

^d Centroid Cp₁, zirconium atom and Centroid Cp₂ angle. In degrees.

^e Accessible surface area Å².

^f Global softness in eV⁻¹.

each particular point in the grid. The probe atom selected was an sp³ C atom with a -1 point charge. This atom corresponds to atom C.3 in the Tripos Force Field [17] which was used to calculate Van der Waals (steric) and Coulomb (electrostatic) interactions.

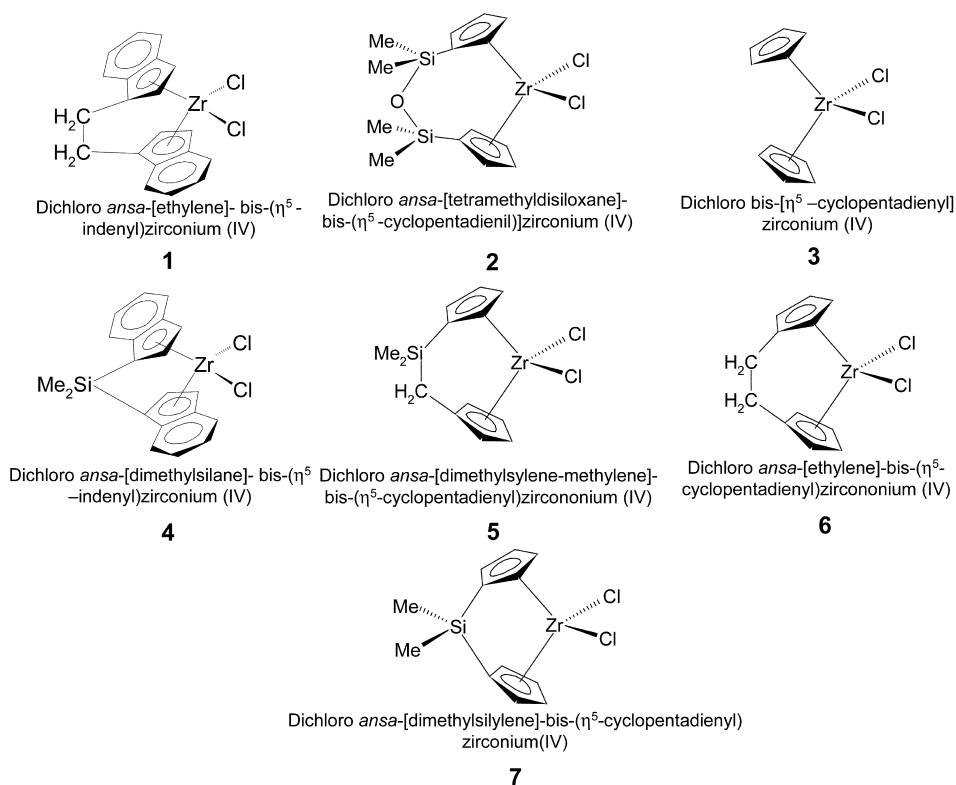
The so-called local softness is another 3D field used to evaluate catalyst reactivity. This field takes into consideration the change in electron density of the cationic active species under the influence of an incoming reagent. The local softness was defined by Yang and Parr as [18]:

$$s(\mathbf{r}) = \left(\frac{\delta\rho(\mathbf{r})}{\delta\mu} \right)_{v(\mathbf{r})} \quad (1)$$

where ρ is the electron density, μ is the chemical potential and $v(\mathbf{r})$ is the external potential. Geerlings [19] and Harbola et al. [20] suggested that local softness can be used as an intermolecular reactivity index, since it can be considered as the global softness index weighted by the Fukui function over the molecule such that it is a measure of the softness density given by the equation:

$$s(\mathbf{r}) = f(\mathbf{r})S \quad (2)$$

where $f(\mathbf{r})$ is the Fukui function defined by Parr and Yang as



Scheme 1.

[21]:

$$f(\mathbf{r}) = \left(\frac{\delta\rho(\mathbf{r})}{\delta N} \right)_{v(\mathbf{r})} \quad (3)$$

Due to the discontinuity of the first derivative with respect to the number of electrons, N , in Eq. (3), the following three functions have to be defined in a finite difference approximation given as:

$$f^+(\mathbf{r}) \approx \rho_{N_0+1} - \rho_{N_0} \quad (4)$$

$$f^-(\mathbf{r}) \approx \rho_{N_0} - \rho_{N_0-1} \quad (5)$$

$$f^0(\mathbf{r}) \approx 1/2(\rho_{N_0+1} - \rho_{N_0-1}) \quad (6)$$

where ρ_{N_0} , ρ_{N_0+1} and ρ_{N_0-1} are the electron density of the system N , $N + 1$ and $N - 1$ electrons. These three indices can be calculated for a nucleophilic, electrophilic and radical reaction, respectively. In our study, the f^+ index for a nucleophilic attack of the monomer on the cationic species of the catalyst is of special importance.

In a similar way, the global softness, S , can be calculated using the following approximation:

$$S = \left(\frac{\delta N}{\delta \mu} \right)_{v(\mathbf{r})} \approx \frac{1}{IE - EA} \quad (7)$$

where IE and EA are the ionisation energy and electron affinity, respectively [21]. Local softness was qualified as the natural DF concept for characterizing a reactive site [19]. LUMO and local softness can be considered useful representations of the electron density in the analysis of nucleophilic reactions such as olefin polymerisation catalysed by metallocenes.

2.3. Alignment rule

To compare all the cationic active species of the catalysts tested, it is essential in 3D-QSAR to align all structures in a common framework [22]. In the present study, the catalyst molecules were aligned such that the active site presented similar orientations. For the molecular alignment represented in Fig. 1, we used the zirconium atom, the cyclopentadienyl centroids Cp_1 and Cp_2 , and the alkyl carbon atom C attached to the metal. The cubic region ($16 \text{ \AA} \times 16 \text{ \AA} \times 16 \text{ \AA}$) used to calculate the 3D fields is also indicated. Several grid spacings were tested, but best results were obtained using a lattice of 1.0 \AA spacing. This value represents a balance between greater precision in the 3D field indicated by a finer grid and the so-called 'brown noise' related to the sensitivity of the statistical technique applied to generate the models [23].

2.4. CoMFA calculations and PLS analysis

We used the CoMFA module [24] implemented in the Sybyl package [25] for all 3D-QSAR analyses.

The 3D fields mentioned above (steric, electrostatic,

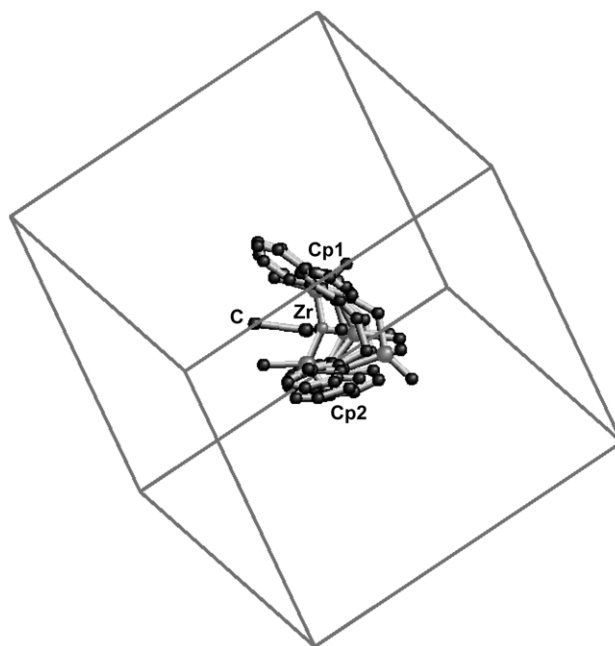


Fig. 1. Catalyst alignments and the cubic region used in the CoMFA calculations.

electron density, HOMO, LUMO and local softness) were evaluated as descriptors able to explain polymerisation activity and polymer molecular weight in the QSAR analysis.

PLS [10] analyses were performed for different combinations of field descriptors. Standard CoMFA scaling was selected on all data matrices to give equal weighting to all fields considered. Leave one out (LOO) [26] cross-validated PLS analysis was initially performed to determine both the robustness of the models and the optimal number of components. This can be achieved by examining the predictive residual sum of squares (PRESS) and the cross-validated regression coefficient (q^2) as guidelines. The q^2 statistics is defined as:

$$q^2 = 1 - \text{PRESS}/\text{SD} \quad (8)$$

where PRESS is given by the sum of the squared deviations of the actual and predicted dependent variables, and SD is the sum of the squared deviations of each dependent variable from the mean of all dependent variables. It has been estimated by some authors [27], that a q^2 value greater than 0.3 has a 95% confidence limit. In drug design, it is common practice to consider helpful a model with a q^2 greater than 0.5, i.e. half way between perfect predictions ($q^2 = 1.0$) and no model at all ($q^2 = 0.0$). The optimum number of components was determined by minimum PRESS and maximum q^2 values. Whenever the increase in q^2 with an additional component was less than 5%, the model with fewest components was selected. Subsequent non-cross-validated PLS analyses were carried out for the optimum number of components to obtain a final model.

The CoMFA results are graphically represented as a 3D

map. For the sake of clarity, only one catalyst from the set is shown in the corresponding figure.

3. Results and discussion

The first step in the search for a quantitative relationship between polymerisation activity, polymer molecular weight and catalyst structure involved identifying simple correlations between these experimental data and scalar descriptors of the catalyst molecular structure. We used geometrical variables that could describe accessibility to the active site to search for such correlation. The steric scalar parameters analysed in this step were the Connolly accessible surface area (ASA) [28] of the metal centre, and the catalyst Cp-Zr-Cp angle. No correlation between the calculated ASA of the metal atom and the experimental data was found. However, good correlation between experimental polymerisation activity and the Cp-Zr-Cp angle was noted for all the catalysts except $[\text{Et}(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ which seems to be an ‘outlier’ (see point 1 in Fig. 2(a)). The behaviour of this catalyst as an outlier is further discussed

below (Section 3.3). Additionally, good correlation was observed between polymer molecular weight and the Cp-Zr-Cp angle, as shown in Fig. 2(b).

For clarity, the results were divided according to the different experimental data, i.e. polymerisation activity and polymer molecular weight. These two sets of results were further divided into subsections corresponding to the fields successfully correlated with the experimental data, i.e. the steric, LUMO and local softness fields.

Best predictive statistics were yielded by models comprised of only one field. Models that considered field combinations gave slightly worse results, probably because of the different nature of each 3D field.

3.1. Polymerisation activity

3.1.1. Steric field

Taking into account the good correlations observed between the scalar geometrical parameters mentioned above and the experimental data, the steric field was considered an appropriate 3D descriptor to explain differences in both activity and molecular weight data. For the polymerisation activity data, poor correlation was found with the steric field, resulting in a cross-validated regression coefficient of $q^2 = 0.102$. Predicted activities for all catalysts were calculated from the cross-validation run and were plotted against the experimental polymerisation activity data. This kind of plot is used in drug design to identify possible outliers when the regression coefficient is too small. In our case, the actual versus predicted activity plot clearly showed the $[\text{Et}(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ compound is an outlier (see Fig. 3). In the absence of the outlier molecule, further PLS computations gave rise to a useful model with a q^2 of 0.527 and a standard error of prediction 4.4×10^6 (g PE/Mol M \times h \times atm) at the two components model. The final non-cross-validated model with two components gave a regression coefficient $r^2 = 0.983$ and a corresponding standard error of estimate (SEE) 0.7×10^6 (g PE/Mol M \times h \times atm).

The standard deviation (stdev) times of the QSAR coefficient (β) fields give a rough location from which structure-activity relationship statements can be inferred, discriminating areas where interactions are important from those that have no significance [9]. Fig. 4 shows isosurfaces for the $\text{stdev}^* \beta$ field contoured at 0.1 (dark grey, positive values) and -0.1 (light grey, negative values) over the most active molecule of the set. Positive values indicate areas where an increase in steric hindrance would produce higher polymerisation activity, and areas showing negative values are those where release of steric bulk would improve catalytic activity. As shown in Fig. 4, putting bulky substituents on the ligands around the active site would result in enhanced catalytic activity. It could be that bulky substituents prevent the anionic co-catalyst species approaching the metal and thus aid the complexing and insertion of the ethylene monomer.

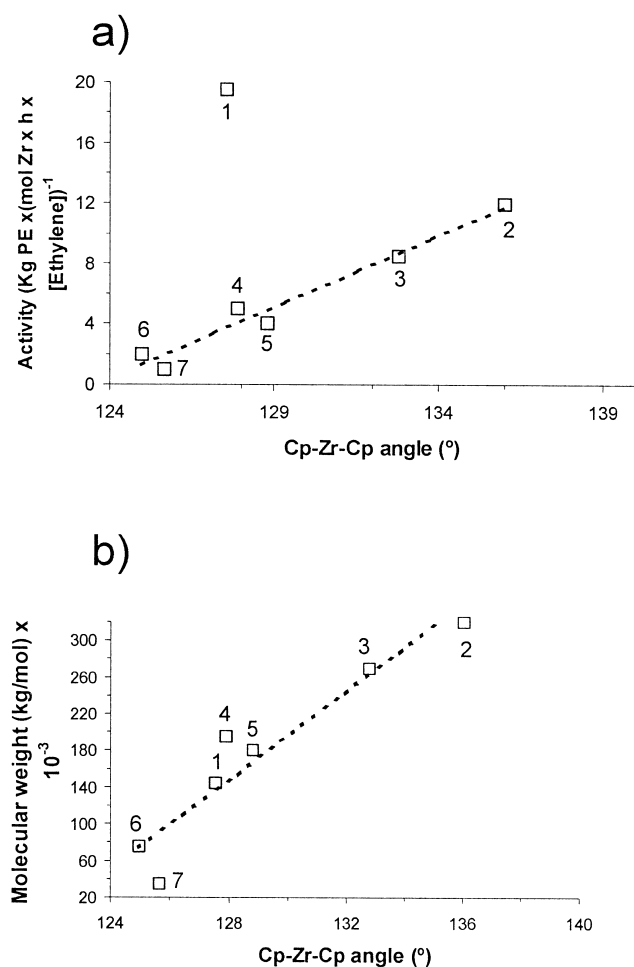


Fig. 2. (a) Plot of polymerisation activity versus Cp-Zr-Cp angle. (b) Plot of polymer molecular weight versus Cp-Zr-Cp angle.

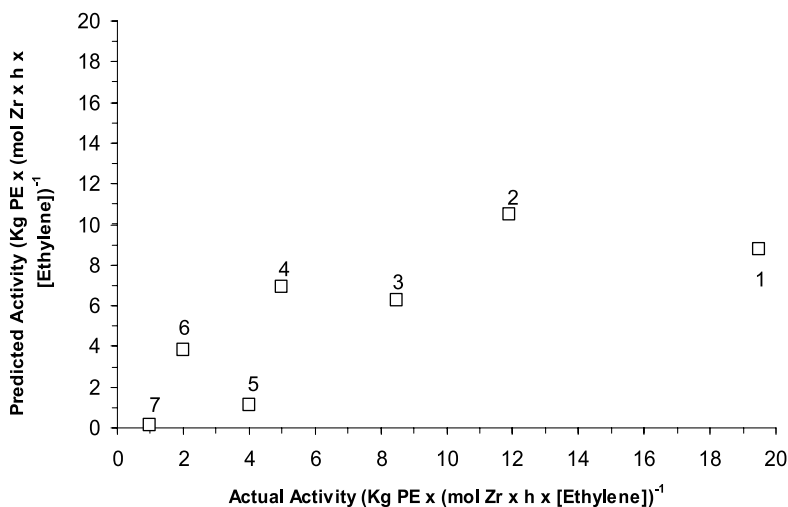


Fig. 3. Residual plot of data (predicted versus actual polymerisation activity data) obtained from the steric field CoMFA model.

3.1.2. LUMO field

3D fields descriptors associated with the electronic structure of the catalysts, such as LUMO and local softness, were found to satisfactorily correlate with polymerisation activity.

For the LUMO, the two components model yielded a q^2 value of 0.643 and a SEP of 5.5×10^6 (g PE/Mol $M \times h \times atm$). These results were obtained when the $[Et(Ind)_2Zr(CH_3)]^+$ outlier was omitted. Similarly, as occurred with the steric field, the outlier was identified in a previous cross-validated PLS run through the corresponding actual/predicted activity plot. The subsequent non-cross-validated run gave an $r^2 = 0.996$ and a $SEE = 0.5 \times 10^6$ (g PE/Mol $M \times h \times atm$). Fig. 5(a) shows isosurfaces for the $stddev^*\beta$ field contoured at 0.1 (dark grey) and -0.1 (light grey), values around the most active catalyst (2) of the set. The above results can be interpreted by plotting the LUMO field corresponding to the same molecule (Fig. 5(b)) contoured at 0.03 (dark grey) and -0.03 (light grey) values. According to the frontier molecular orbital theory, the LUMO corresponds to molecular regions where the addition of an electron is

energetically more favourable. Olefin polymerisation is considered to be a nucleophilic reaction, so that the incoming ethylene monomer will attack the catalyst species at the most favourable area for electron addition (light grey areas in Fig. 5(b)). In the case of the metallocene catalytic active species, the LUMO is formed by two phases: a positive one surrounding the metal atom and a negative phase with two lobes pointing towards the so-called frontside and backside positions of the catalyst [29]. The $stddev^*\beta$ field indicates that an increase in the LUMO contribution to the frontside lobe (negative phase, light grey contours in Fig. 5) and to the LUMO around the zirconium atom (positive phase, dark grey contour in Fig. 5) should improve the polymerisation activity of the catalyst. To find out which structural changes might be suitable for increasing catalytic activity, the effect of the molecular structure of the catalyst on the LUMO shape needs to be

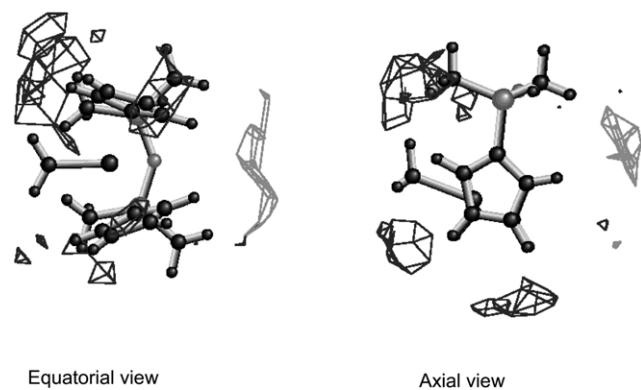


Fig. 4. Standard deviation time coefficient CoMFA maps showing areas of favourable and unfavourable steric interaction in terms of polymerisation activity.

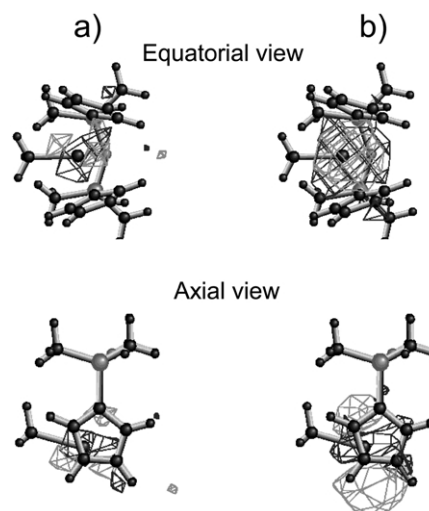


Fig. 5. (a) Standard deviation time coefficient CoMFA maps showing areas of favourable and unfavourable LUMO field contributions for polymerisation activity case. (b) LUMO field contoured at the levels 0.03 au (dark grey) and -0.03 au (light grey).

considered. The LUMO is formed by atomic orbitals (AOs) belonging to the metal, as well as to some of the aromatic ligand's atoms. To clarify the influence of the fused aromatic substituent attached to the Cp ligand (i.e., indenyl and fluorenyl ligands) on LUMO shape, we calculated the difference between indenyl and cyclopentadienyl based compounds. Fig. 7 shows the LUMO of the silyl-bridged bisindenyl zirconocene (**4**) minus the LUMO of the silyl-bridged biscyclopentadienyl zirconocene (**7**). The most significant difference is that the atoms located in the indenyl ligand around the active site contribute towards enhancing the LUMO on the frontside of the catalyst.

On the contrary, we observed no contribution of the atoms of the bridge moiety to the LUMO. To assess the effect of the bridge on the resulting LUMO, the difference between the LUMOs of two distinct molecules with and without a bridge was calculated. Fig. 6 shows the contour obtained by subtracting the LUMO of the ethyl-bridged bis-Cp catalyst molecule (**6**) from the LUMO of the unbridged bis-Cp zirconocene (**3**). The positive contour (dark grey lines) indicates areas where LUMO values are larger for the unbridged (**3**) than for the bridged zirconocene (**6**). The negative contour (light grey) corresponds to LUMO contributions which are larger for the bridged than for the unbridged molecules. It can be observed that the effect of the ethyl bridge (**6**) is to close the Cp-Zr-Cp angle enhancing then the participation of the Cp ring atoms near the bridge and far from the active site in the LUMO. An increased Cp-Zr-Cp angle such as that of the unbridged molecule (**3**) allows a larger contribution to the LUMO of the Cp atoms near the active centre. Therefore, the role of the bridge is to modify the alignment of the Cp ligands with respect to the metal centre so that atomic orbitals belonging to the ligand atoms can favourably interact with AOs of the Zr atom, thus increasing catalytic activity.

In conclusion, the above observations clearly indicate that AO contributions to the LUMO of atoms in the vicinity of the active site increase the polymerisation activity. The atomic orbitals of the ligand atoms need to be appropriated aligned to effectively overlap the orbitals of the metal atom. This can be achieved by increasing the Cp-Zr-Cp angle (at

least for low steric demanding monomers, such as ethylene) and/or using bulky aromatic ligands, i.e. indenyl or fluorenyl.

3.1.3. Local softness

A significant correlation between the local softness field and the polymerisation activity was found. Using this descriptor, the predicted versus actual activity data plot again reveals that the $[\text{Et}(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ (**1**) catalyst is an outlier. The cross-validated q^2 value calculated after removing the outlier was 0.573 at the two components model with a standard error of prediction of 3.9×10^6 (g PE/Mol M \times h \times atm). The final non-cross-validated model with two components gave a regression coefficient $r^2 = 0.986$ with an SEE = 0.6×10^6 (g PE/Mol M \times h \times atm). Fig. 8(a) shows isosurfaces of the $\text{stdev}^*\beta$ field contoured over the most active catalyst (**2**). It can be seen that an increase in the local softness around the metal centre is required to enhance catalytic activity. The general shape of the local softness is shown in Fig. 8(b). As in the LUMO representation, two lobes are oriented towards the frontside and backside positions along with an area located around the metal atom. In this case, all these regions have positive values in contrast with the two different phases (positive and negative) described for the LUMO descriptor. The atoms in the bridge show no significant contribution to the local softness. However, the indirect influence of the bridge on the local softness distribution needs to be clarified. Fig. 9 shows the local softness isosurface at a 0.01 level for four bis-Cp catalysts (**2,3,6,7**) considered in this paper. The two most active catalysts (**2,3**) (see Fig. 9(a) and (b)) show larger local softness between the Cp ligands and the metal centre compared to the other less active bis-Cp compounds (**6,7**) (see Fig. 9(c) and (d)). It seems, that the local softness is best shared between the metal atom and the Cp ligand when the corresponding atoms overlap more effectively. In addition, when comparing Cp- versus Ind-based catalysts, it was found that the whole indenyl ring contributes more to local softness than the Cp ring around the metal atom. As a consequence, catalyst (**4**) is more active than catalyst (**7**).

From the arguments above, it can be concluded than

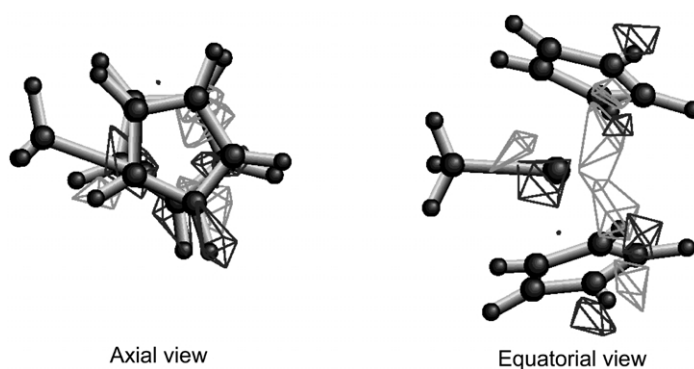


Fig. 6. Difference between the unbridged zirconocene (**3**) and the ethyl-bridged bis-Cp catalyst (**6**) LUMO fields contoured at 0.005 au (dark grey) and -0.005 au (light grey).

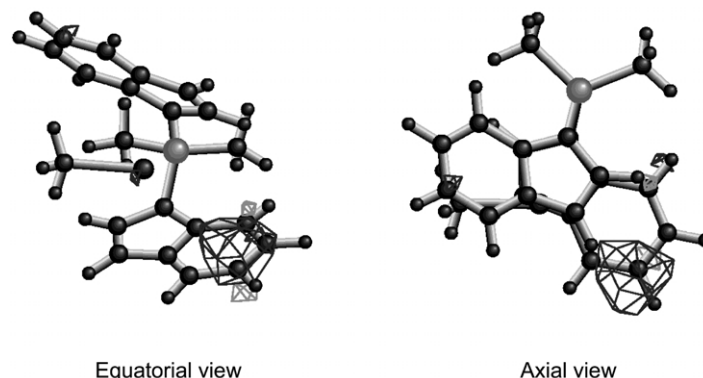


Fig. 7. Difference between the silyl-bridged bisindenyl zirconocene (**4**) and the silyl-bridged biscyclopentadienyl zirconocene (**7**) LUMO fields contoured at 0.003 au (dark grey) and -0.003 au (light grey).

electronic effects between the metal atom and the aromatic ligands are highly influential for polymerisation activity. The alignment of the aromatic ligands with respect to the metal atom centre affects the catalytic activity of the active species, and the bridge only determines how the aromatic ligands can overlap with the metal centre in a more or less effective way.

As discussed, it seems to be that the local softness and LUMO fields contain similar information about the electronic structure of the catalysts and cannot therefore be used in conjunction in QSAR studies. To analyse the degree of correlation between these fields (local softness and LUMO), the Tanimoto similarity index [30] was calculated. The Tanimoto index is defined as:

$$T(x, y) = \frac{\sum X_i Y_i}{\sum X_i^2 + \sum Y_i^2 - \sum X_i Y_i} \quad (9)$$

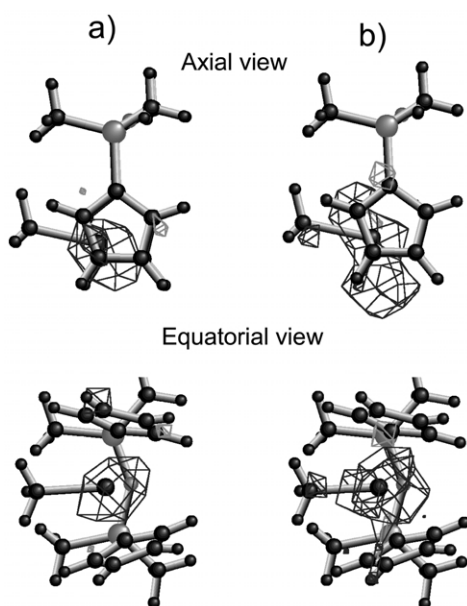


Fig. 8. (a) Standard deviation time coefficient CoMFA maps showing areas of favourable and unfavourable local softness field contributions to polymerisation activity. (b) Local softness field contoured at 0.02 au (dark grey) and -0.02 au (light grey).

where X_i and Y_i are the field values at the i grid point. It should be noticed that in order to calculate the Tanimoto index, the values of the LUMO at each grid point were scaled up according to the mean value of local softness and LUMO.

Similar fields have a $T(x, y)$ coefficient close to 1.0, whereas unrelated fields are close to 0.0. The Tanimoto index calculated for the compounds examined here are within the range 0.3–0.5, indicating no significant correlation among these electronic fields for all the catalysts.

3.2. Polymer molecular weight

3.2.1. Steric field

Similarly to the case of polymerisation activity, correlations between the polymer molecular weight and steric field were also performed. The results were not too conclusive. Nevertheless, the MW and the Cp-Zr-Cp angle showed better linear correlation than in the case of

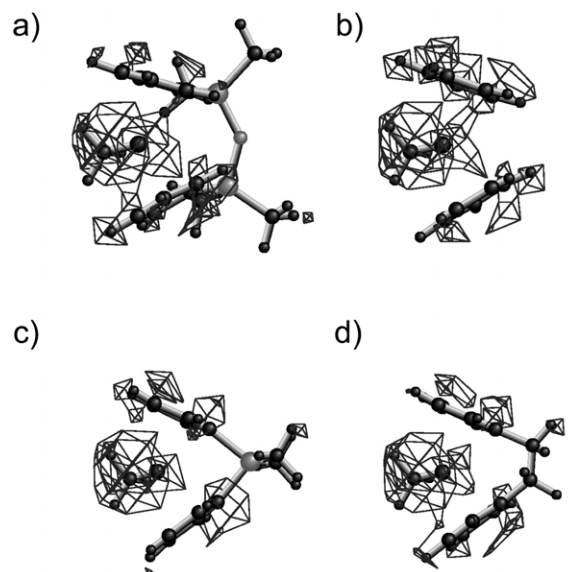


Fig. 9. Equatorial view of the local softness field contoured at 0.01 au for the catalysts: (a) catalyst number **2**, (b) catalyst number **3**, (c) catalyst number **6**, (d) catalyst number **7**.

polymerisation activity (Fig. 2(b)). Furthermore, no outlier was identified based on polymer molecular weight.

The lack of correlation between polymer molecular weight and steric field can be explained by the fact that the rotational energy barriers for the β -agostic conformers, obtained by means of dihedral angle scans around the $C\alpha-C\beta$ bond, are very low (1–2 kcal/mol) for all catalysts. These results support the idea that according to the most accepted β -transfer mechanisms [31], ligand size has a limited influence on polymer molecular weight, at least for the species considered here.

3.2.2. LUMO field

PLS analyses revealed satisfactory correlation of the LUMO with polymer MW. The cross-validated q^2 value was 0.714 and the SEP = 74.3×10^3 g/mol at the two components model, whereas the values for the fitted model were $r^2 = 0.990$, and SEE = 14.2×10^3 g/mol. In Fig. 10, the $\text{stdev}^*\beta$ field shows similar features to those described for polymerisation activity, i.e. a positive contribution to the positive phase around the metal atom and a negative contribution to the negative phase of the LUMO in a frontside position. These properties should lead to increased polymer molecular weight. In addition to these common features observed for polymerisation activity and polymer molecular weight, we noted a contribution to the LUMO positive phase by atoms belonging to the indenyl ligands. This should also help to obtain higher polymer molecular weights.

3.2.3. Local softness field

Correlating the MW with local softness gave a q^2 value of 0.679 for two components and a SEP of 77.6×10^3 g/mol. We obtained a correlation coefficient $r^2 = 0.997$ and SEE = 7.8×10^3 g/mol in the non-cross-validated run for the optimal number of components. Fig. 11 shows the $\text{stdev}^*\beta$ field contoured on the most active catalyst (2). The plotted isosurface is very similar to that obtained in the polymerisation activity data analysis. The

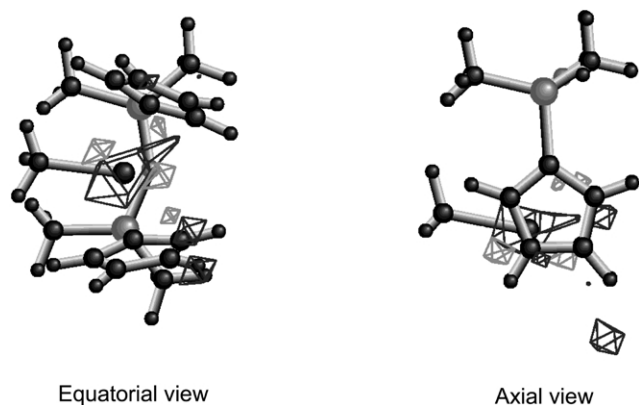


Fig. 10. Standard deviation time coefficient CoMFA maps showing areas of favourable and unfavourable LUMO field contributions to polymer molecular weight.

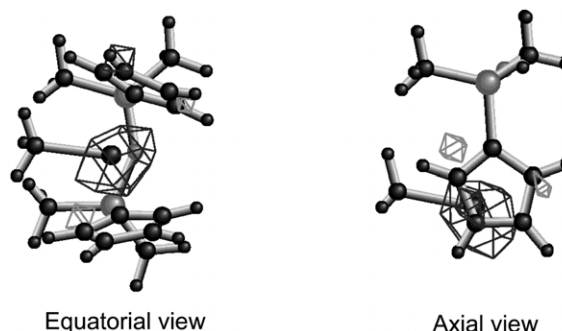


Fig. 11. Standard deviation time coefficient CoMFA maps showing areas of favourable and unfavourable local softness field contributions to polymer molecular weight.

dark grey areas show that a positive contribution to the local softness around the metal atom will lead to an increased polymer molecular weight. Further, in this model more emphasis is given to areas on the aromatic ligands indicating that an increased local softness in these regions should help obtain higher molecular weight polyethylene. On the contrary, the light grey contour in Fig. 11 suggests that local softness in the vicinity of the bridge has a detrimental effect on polymer molecular weight.

3.3. Predictive capacity of CoMFA models

The main benefits of computer aided molecular design in general and particularly when applied to the design of polymerisation catalysts, is to produce a rank list of the best catalyst candidates before synthesis. Hence, CoMFA could be a very useful tool for designing new catalysts with the desired properties. The underlying PLS technique is a QSAR method suitable for predicting the performance of a catalyst, rather than only reproducing experimental results based on the features of the catalyst.

Models obtained from PLS calculations have two main applications. On one hand, they provide useful information about which changes should be made to the catalyst molecules to enhance its catalytic activity or to increase polymer molecular weight. This information has been described in the foregoing sections.

On the other hand, CoMFA models can also be used to predict the catalytic activity and the polymer molecular weight that could be expected for any new catalyst, provided it does not differ too much from the catalysts used in the training set.

Two cases were worked out to show the predictive capacity of the models developed here. In the first, the catalyst activity of the $[\text{Et}(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ compound was recalculated using the present model based on the LUMO field alone, which gives the best predictive performance (see Section 3.1.3). This catalyst was identified as an outlier in all the QSAR models of polymerisation activity data. To explain this behaviour, we consulted the available experimental data. Thus, Kaminsky [32] reported polymerisation

activity and molecular weight data for some of the molecules considered in our study, including the outlier catalyst. Fig. 12(a) and (b) show our polymerisation activity and molecular weight experimental values versus those of Kaminsky. Although the experimental data were obtained under different polymerisation conditions, the two molecular weight datasets correlate well. In contrast, there is clear discrepancy between activity measurements for the $[\text{Et}(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ compound, confirming the outlier behaviour of this catalyst (1).

Using the LUMO CoMFA model, we calculated the predicted polymerisation activity for this compound at 6.7×10^6 g PE/(mol M \times h \times atm). This value satisfactorily fits the plot correlating our experimental data with Kaminsky's data (Fig. 12(a)) reflecting the predictive capacity of this model.

In the second case, polymerisation activity and polymer molecular weight were calculated for the unbridged $[(\text{Ind})_2\text{Zr}(\text{CH}_3)]^+$ catalysts using the LUMO CoMFA model. This compound, which was not included in the original training set, has two structural characteristics required for high catalytic activity and polymer

molecular weight, i.e. an open Cp-Zr-Cp angle (141°) and a fused aromatic substituent attached to the Cp (indenyl group), which contributes considerably to the LUMO, as discussed above. The predicted polymerisation activity for this catalyst was 18.8×10^6 g PE/(mol M \times h \times atm) and the predicted polymer molecular weight was 415×10^3 g/mol. Alt et al. [1] performed a comparative study on several catalysts including polymerisation activity and polymer molecular weight data for the unbridged bis-indenyl and bis-cyclopentadienyl metallocene catalysts. Table 2 lists their results alongside our predictions for the bis-indenyl catalyst. The experimentally observed tendencies are well captured by the model despite the different polymerisation conditions used by these authors. The bis-indenyl compound is twice as active as the bis-cyclopentadienyl catalysts in both the experimental and predicted values. Similarly, polymer molecular weight is about 1.5 times higher for the indenyl compared to the Cp-based catalyst. For this particular compound, the model behaved satisfactorily despite predicted values being well above the data used in the training set.

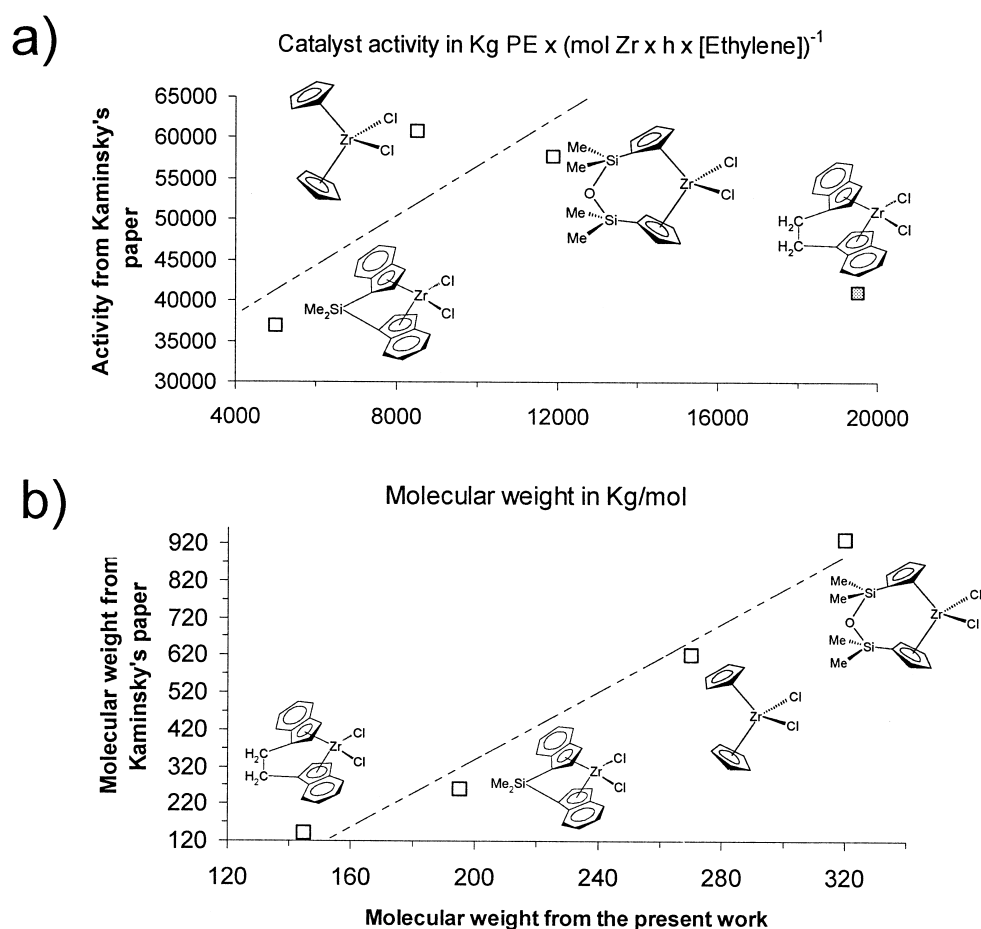




Fig. 12. Comparison between the present experimental data and those reported by Kaminsky [32]. (a) Polymerisation activity (outlier indicated by a shaded square). (b) Polymer molecular weight.

Table 2
Predictive capability of the model

	Experimental activity ^{a,b}	Predicted activity ^c	Experimental MW ^{a,d}	Predicted MW ^d
	1490	8500	290	270
	3200	18,800	470	415

As can be noted, activity units from Alt's paper are different to activity units considered in our model. This is the reason of the large deviation between the Alt's activities and those reported here. It should be pointed out that the polymerisation reaction conditions are also different in both works. For details in this topic see the computational methods section and Ref. [1].

^a From Ref. [1].

^b In kg PE (g Zr × h)⁻¹.

^c In kg PE (mol Zr × h × [Ethylene])⁻¹.

^d In kg/mol.

4. Conclusions

To our knowledge, this is the first attempt made to correlate experimental data such as polymerisation activity and polymer molecular weight with the 3D structural properties of catalyst active species. The models used were obtained through the well known 3D-QSAR method successfully used in drug design. In this type of analysis, the experiments have to be performed in the same conditions for each catalyst. The aim is to separate the contributions related to the structure of the organometallic complex from other effects. We used a set of metallocene catalysts and determined polymerisation activity and polymer molecular weight under controlled conditions. The statistical robustness of the resulting model, measured mainly through cross-validation, suggests further work should be conducted using more catalysts.

The initial assumption that the active species is the cationic complex seems to be adequate as a common structural framework for applying 3D-QSAR procedures.

Variation in polymerisation activity data can be successfully explained in terms of steric and electronic fields. The calculated model predicts that an increase in the Cp-Zr-Cp angle and/or incorporation of bulky ligands will enhance catalytic activity.

The effect of electronic interaction was confirmed by correlations found between activity and the LUMO molecular orbital and between activity and local softness. The model revealed that the arrangement of the aromatic ligands around the metal centre as well as the chemical nature of the ligand (i.e., Cp or Ind) significantly contribute to explaining the variance shown by the experimental data. The structure of the bridge is not directly implicated in these two fields (LUMO and local softness) except that it forces the Cp-Zr-Cp angle and thus to a greater or lesser extent promotes electronic interaction between the metal centre and the atoms of the ligands.

Polymer molecular weight was found to correlate well

with the LUMO and local softness fields. The Cp-Metal-Cp angle is the key geometric variable enabling a particular distribution in the LUMO and local softness field that lead to increased polymer molecular weight. Variation in the MW dataset cannot be satisfactorily explained by the steric field.

The information extracted from the model evokes numerous possible structural modifications that could be used in the design of tailor-made catalysts.

Acknowledgements

Thanks are due to the MICYT (Grant MAT2002-01242) for financial support. The authors also thank the CESGA (Centro de Supercomputacion de Galicia) for the use of their computational resources.

References

- [1] Alt HG, Koppl A. *Chem Rev* 2000;100:1205–21.
- [2] Martin YC. *Quantitative drug design*. New York: Marcel Dekker; 1978.
- [3] Hansch C, Klein T. *Acc Chem Res* 1986;19:392.
- [4] Möring PC. *J Mol Catalysis* 1992;77:41.
- [5] Yao S, Shoji T, Iwamoto Y, Kamei E. *Comp Theor Polym Sci* 1999;9: 41–6.
- [6] Linnolahti M, Pakkanen TA. *Macromolecules* 2000;33:9205–14.
- [7] Green SM, Marshall GR. *Trends Pharmacol Sci* 1995;16:285.
- [8] Kim K. *J Comput-Aided Mol Design* 1995;9:308.
- [9] Cramer III RD, DePriest SA, Patterson DE, Hecht P. In: Kubinyi H, editor. *The developing practice of comparative molecular field analysis in 3D QSAR in drug design: theory, methods and applications*. Dordrecht: Kluwer; 2000. p. 443–85.
- [10] Wold S, Ruhe A, Wold H, Dunn WJ. *SIAM J Sci Stat Comput* 1984;5: 735–43.
- [11] Cruciani G, Baroni M, Clementi S, Costantino G, Riganelli D, Skagerberg B. *J Chemometrics* 1992;6:335–46.
- [12] Martin YC, Kim KH, Liu CT. *Comparative molecular field analysis: CoMFA*. *Advances in Quant. Str.-Property Relationships*, vol. 1. JAI Press, Greenwich CT; 1996. pp. 1–52.

- [13] Froese RDI, Musaev DG, Morokuma K. *J Am Chem Soc* 1998;120:1581.
- [14] Hay PJ, Wadt WR. *J Chem Phys* 1985;82:299.
- [15] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann E, Burant JC, Dapprich S, Millan JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuk AD, Raghavachari K, Foresman JB, Ciolowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith TA, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA. *Gaussian98 (Revision A.1)*, Gaussian Inc., Pittsburgh, PA; 1998.
- [16] Breneman CM, Wiberg KB. *J Comp Chem* 1990;11:361.
- [17] Clark M, Cramer III RD, Van Opdenbosch N. *J Comp Chem* 1989;10:982–1012.
- [18] Yang W, Parr RG. *Proc Natl Acad Sci USA* 1985;82:6723.
- [19] Geerlings P, de Proft F, Langenaeker W. In: Seminario J, editor. *Density functional theory: a source of chemical concepts and a cost-effective methodology for their calculation*. Invited contribution to advances in quantum chemistry, vol. 33. New York: Academic Press; 1999. p. 303. Chapter 17.
- [20] Harbola MK, Chattaraj PK, Parr RG. *Isr J Chem* 1991;35:395.
- [21] Parr RG, Yang W. *Density functional theory of atoms and molecules*. New York, NY: Oxford University Press; 1989.
- [22] Martin YC, Kim KH, Liu CT. *Comparative Molecular Field Analysis: CoMFA*. *Advances in Quant. Str.-Property Relationships*, vol. 1. JAI Press; 1996. p. 1–52.
- [23] Rännar P, Lindgren S. *Chemometrics* 1994;8:111–25.
- [24] Cramer III RD, Patterson DE, Bunce JD. *J Am Chem Soc* 1988;110:5959–67.
- [25] SYBYL, *Molecular Modeling System*, Tripos Inc., 1699 S. Hanley Road, St Louis, MO, 63144.
- [26] Cramer III RD, Bunce JD, Patterson DE, Frank IE. *Quant Struct-Act Relat* 1988;7:18–25.
- [27] Clark M, Cramer III RD. *Quant Struct-Act Relat* 1993;12:137–45.
- [28] Connolly ML. *Science* 1983;221:709–13.
- [29] Muñoz-Escalona A, Ramos J, Cruz V, Martínez-Salazar J. *J Polym Sci, Part A: Polym Chem* 2000;38:571–82.
- [30] Willet PJ. *Similarity and clustering in chemical information systems*. Letchworth: Research Studies Press; 1987.
- [31] Kawamura-Kuribayashi H, Koga N, Morokuma K. *J Am Chem Soc* 1992;114:8687.
- [32] Kaminsky W. *Macromol Chem Phys* 1996;197:3907–45.